Subgradient

Nipun Batra

IIT Gandhinagar

July 31, 2025

Outline

The Problem: Non-Differentiable Functions

Challenge: What happens when we can't compute gradients?

The Problem: Non-Differentiable Functions

Challenge: What happens when we can't compute gradients?

Common Examples in ML

The Problem: Non-Differentiable Functions

Challenge: What happens when we can't compute gradients?

Common Examples in ML

- **L1 regularization:** $f(x) = |x|$ (not differentiable at $x = 0$)

The Problem: Non-Differentiable Functions

Challenge: What happens when we can't compute gradients?

Common Examples in ML

- **L1 regularization:** $f(x) = |x|$ (not differentiable at $x = 0$)
- **Hinge loss:** $f(x) = \max(0, 1 - x)$ (kink at $x = 1$)

The Problem: Non-Differentiable Functions

Challenge: What happens when we can't compute gradients?

Common Examples in ML

- **L1 regularization:** $f(x) = |x|$ (not differentiable at $x = 0$)
- **Hinge loss:** $f(x) = \max(0, 1 - x)$ (kink at $x = 1$)
- **ReLU activation:** $f(x) = \max(0, x)$ (corner at $x = 0$)

The Problem: Non-Differentiable Functions

Challenge: What happens when we can't compute gradients?

Common Examples in ML

- **L1 regularization:** $f(x) = |x|$ (not differentiable at $x = 0$)
- **Hinge loss:** $f(x) = \max(0, 1 - x)$ (kink at $x = 1$)
- **ReLU activation:** $f(x) = \max(0, x)$ (corner at $x = 0$)

The Problem: Non-Differentiable Functions

Challenge: What happens when we can't compute gradients?

Common Examples in ML

- **L1 regularization:** $f(x) = |x|$ (not differentiable at $x = 0$)
- **Hinge loss:** $f(x) = \max(0, 1 - x)$ (kink at $x = 1$)
- **ReLU activation:** $f(x) = \max(0, x)$ (corner at $x = 0$)

Solution: Subgradients

Subgradients generalize gradients to convex but non-differentiable functions

Subgradient

Key Insight: Even when derivatives don't exist, we can still do optimization!

Subgradient

Key Insight: Even when derivatives don't exist, we can still do optimization!

What is a Subgradient?

For convex function f at point x_0, vector g is a **subgradient** if:

$$f(x) \geq f(x_0) + g^T(x - x_0) \quad \forall x$$

Subgradient

Key Insight: Even when derivatives don't exist, we can still do optimization!

What is a Subgradient?

For convex function f at point x_0, vector g is a **subgradient** if:

$$f(x) \geq f(x_0) + g^T(x - x_0) \quad \forall x$$

Subgradient

Key Insight: Even when derivatives don't exist, we can still do optimization!

What is a Subgradient?

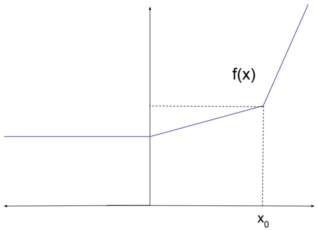
For convex function f at point x_0, vector g is a **subgradient** if:

$$f(x) \geq f(x_0) + g^T(x - x_0) \quad \forall x$$

Intuition: g defines a linear lower bound that touches f at x_0

The Classic Example: $f(x) = |x|$

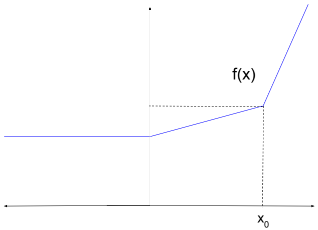
Challenge: Find "derivative" of $f(x) = |x|$ at $x = 0$



Non-differentiable function at $x = 0$

The Classic Example: $f(x) = |x|$

Challenge: Find "derivative" of $f(x) = |x|$ at $x = 0$



Non-differentiable function at $x = 0$

Classical calculus: Derivative doesn't exist at $x = 0$
Subgradient calculus: We can find a set of valid subgradients!

Pop Quiz: Subgradient Intuition

Quick Quiz 1

For $f(x) = |x|$ at $x = 0$, which values could be subgradients?

a) Only $g = 0$

Answer: b) Any value in $[-1, 1]$ satisfies the subgradient condition!

Pop Quiz: Subgradient Intuition

Quick Quiz 1

For $f(x) = |x|$ at $x = 0$, which values could be subgradients?

a) Only $g = 0$
b) Any $g \in [-1, 1]$

Answer: b) Any value in $[-1, 1]$ satisfies the subgradient condition!

Pop Quiz: Subgradient Intuition

Quick Quiz 1

For $f(x) = |x|$ at $x = 0$, which values could be subgradients?

a) Only $g = 0$
b) Any $g \in [-1, 1]$
c) Only $g = 1$ or $g = -1$

Answer: b) Any value in $[-1, 1]$ satisfies the subgradient condition!

Solution

- Construct a differentiable $g(x)$

Solution

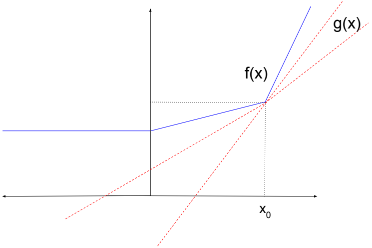
- Construct a differentiable $g(x)$
  - Intersecting $f(x)$ at $x = x_0$

Solution

- Construct a differentiable $g(x)$
  - Intersecting $f(x)$ at $x = x_0$

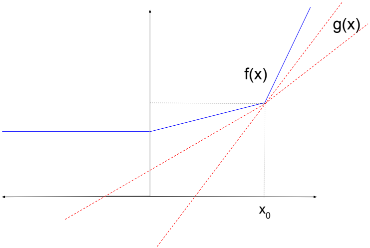
Solution

- Construct a differentiable $g(x)$
  - Intersecting $f(x)$ at $x = x_0$
  - Below or on $f(x)$ for all x



Solution

- Compute slope of $g(x)$ at $x = x_0$



Another Example: $f(x) = |x|$

- Subgradient of $f(x)$ belongs to $[-1, 1]$